AD-

18066

UNCLASSIFIED

ON THE DISTRIBUTION OF QUADRATIC FORMS

by

JAMES PACHARES

Special report to the Office of Naval Research of work at Chapel Hill under Project NR 042 031 for research in probability and statistics.

U. of North Carolina
Institute of Statistics
Mimeograph Series No. 75

July 15, 1953

ACKNOWLEDGMENT

The writer wishes to express his thanks to Professors Harold Hotelling and S. N. Roy for their valuable suggestions and criticisms during the preparation of this work.

Acknowledgment is also made to the Office of Naval Research for financial aid.

TABLE OF CONTENTS

INTRODUCTION

One could easily justify the study of the distribution of quadratic forms from the standpoint that many of the tests in statistics are based on the distributions of quantities which can be thought of as special cases of either a quadratic form or functions of quadratic forms. The applications are too numerous to mention; however, we shall list a few as illustrations.

(i) The distribution of a definite quadratic form where the components have a multivariate normal distribution.

(ii) The problem of finding the power function of the chi-square statistic, for large samples, can be reduced to that of finding the distribution of a positive definite quadratic form in non-central normal variates.

(iii) The distribution of a form of the serial correlation coefficient can be expressed in terms of the distribution of a ratio of two quadratic forms. See Anderson [1][1]

(iv) Of special importance is von Neumann's statistic, [19], [20], the ratio of

1. Numbers in square brackets refer to bibliography.

the mean square successive difference to the variance, used to test whether observations are independent or whether a trend exists.

(v) Durbin and Watson [6], use a similar statistic to test the error terms for independence in least squares regression.

(vi) Koopmans [12], says, "Assuming a normal distribution for the random disturbance, the mathematical prerequisite for an estimation theory of stochastic processes is the study of the joint distributions of certain quadratic forms in normal variables". The problem Koopmans considers is that of estimating the serial correlation in a stationary stochastic process.

(vii) To test hypotheses concerning variance components in the analysis of variance, we require the distribution of an indefinite quadratic form.

(viii) Hotelling [9], shows how the distribution of the ratio of an indefinite quadratic form in non-central normal variates to a definite quadratic form could be used in the theory of selecting variates for use in prediction.

(ix) McCarthy [13], shows how the distribution of the ratio of two definite quadratic forms could be used to make an F test in the analysis of variance when the assumptions of equal variances and independence of the observations are not met.

It would make this report quite lengthy to discuss the distributions of all these functions of quadratic forms. We shall restrict ourselves to the study of a definite quadratic form in both central and non-central independent normal variates, giving an important application for each distribution. Then we shall discuss two special cases of an indefinite quadratic form. Finally, we shall discuss a few inequalities. We give below a slightly more detailed chapter-wise breakdown.

In Chapter I we shall be concerned with the distribution of a definite quadratic form in independent $N(0,1)$ variates. Robbins [16], has treated this problem but we have carried it a bit further. Robbins and Pitman [17], have given an expression for the distribution of a linear combination of chi-square variates. We feel that we have improved on this form. We have derived an expression which depends only on the value of the determinant of the form and on the moments of the inverse quadratic form. The expression is an alternating series which converges absolutely and is such that if we stop after any even power we have an upper bound, and if we stop after any odd power, a lower bound to the

cumulative distribution function. Hotelling [10] and Gurland [8], have suggested the use of Laguerre polynomials in finding distributions of quadratic forms. A brief account of Hotelling's method will be given.

In Chapter II we have derived an expression for the distribution of a definite quadratic form in non-central independent normal variates which depends only on the value of the determinant of the form and on the moments of the inverse quadratic form in normal variates with imaginary means. This statement will be made clearer later on. This result enables us to find the power function of the chi-square statistic.

In Chapter III we have discussed the distribution of the difference between two independent chi-squares having different numbers of degrees of freedom. If the degrees of freedom are the same, the distribution becomes the same as the distribution of the sample covariance in sampling from a normal population. We have studied the properties of this distribution in some detail.

In Chapter IV we give some inequalities for the distribution of a quadratic form in $N(0, 1)$ variates and also for the general case.

NOTATION

All vectors are column vectors and primes indicate their transposes.

"p.d.f." stands for "probability density function";

"c.d.f." stands for "cumulative distribution function";

"r.v." stands for "random variable";

"q.f." stands for "Quadratic form" ;

"$N(\mu,\sigma)$" stands for "a r.v. having a normal p.d.f. with mean μ and standard deviation σ".

CHAPTER I

THE DISTRIBUTION OF A DEFINITE QUADRATIC FORM IN INDEPENDENT CENTRAL NORMAL VARIATES

1.1 The problem.

Suppose we have a q.f. $Q_n = \frac{1}{2} Y'A\,Y$ in $Y_1,\ldots,Y_n$, where the Y_i are independent $N(0, 1)$ variates, and where $Y' = (Y_1,\ldots,Y_n)$, where the dash denotes the transpose of the column vector Y. Let $F_n(t) = \Pr(Q_n \leq t)$. Then the problem is to find $F_n(t)$. It is well known that we can make an orthogonal transformation, $Y = P\,X$, say, where $P\,P' = P'P = I$, such that $\frac{1}{2} Y'A\,Y = \frac{1}{2} X'P'APX = \frac{1}{2} X'D_a X = \frac{1}{2} \sum_{i=1}^{n} a_i x_i^2$, where D_a is a diagonal matrix having the elements $a_1,\ldots,a_n$ in its main diagonal, and where $a_1,\ldots,a_n$ are the latent roots of the matrix A. Under such a transformation, $X_1,\ldots,X_n$ remain independent $N(0, 1)$. So the problem is now to find the $\Pr(\frac{1}{2}\sum_{i=1}^{n} a_i X_i^2 \leq t)$, where we assume that $a_i > 0$, $i=1,\ldots,n$.

1.2 The solution.

Theorem 1.1. Let $Q_n = \frac{1}{2}\sum_{i=1}^{n} a_i x_i^2$, where the x_i are independent $N(0, 1)$, and where $a_i > 0$, $i=1,\ldots,n$. Let $Q_n^* = \frac{1}{2}\sum_{i=1}^{n} a_i^{-1} X_i^2$.

Then,

$$\text{(a)}\quad F_n(t) = \frac{t^{n/2}}{(a_1 \ldots a_n)^{1/2}} \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} \frac{E(Q_n^*)^k}{\Gamma(\frac{n}{2}+k+1)}$$

where $E(Q_n^*)^k$ is the k-th moment of Q_n^*.

(b) The series is absolutely convergent and therefore it is convergent.

(c) For any two positive integers r and s and every $t > 0$,

$\sum_{k=0}^{2s-2} d_k > F_n(t) > \sum_{k=0}^{2r-1} d_k$, where

$$d_k = \frac{t^{n/2}}{(a_1 \ldots a_n)^{1/2}} \frac{(-t)^k}{k!} \frac{E(Q_n^*)^k}{\Gamma(\frac{n}{2}+k+1)} \quad .$$

Proof.

Let $d\underline{x} = dx_1 \ldots dx_n$ and $[R] = [\frac{1}{2} \sum_{i=1}^{n} a_i X_i^2 \leq t]$, then

$$F_n(t) = (2\pi)^{-\frac{n}{2}} \int \ldots \int_{[R]} \exp - \frac{1}{2} \sum_{1}^{n} x_i^2 \, d\underline{x}.$$

We shall make use of the Dirichlet integral;

$$\int \cdots \int \prod_{j=1}^{n} x_j^{\ell_j - 1} dx_j = \frac{\prod_{j=1}^{n} \Gamma(\frac{\ell_j}{p_j})\ c_j^{\ell_j}}{\Gamma[\sum_1^n (\frac{\ell_j}{p_j}) + 1]} ,$$

where $-\infty < x_i < \infty$, and ℓ_j, c_j, p_j are all positive, and the range of integration is $\sum_1^n (\frac{x_i}{c_i})^{p_i} \leq 1$. See Edwards [7]. If we expand the exponential in the integrand, we get

$$F_n(t) = (2\pi)^{-\frac{n}{2}} \int_{[R]} \cdots \int \sum_{k=0}^{\infty} \frac{(-\frac{1}{2} \sum_1^n x_i^2)^k}{k!} \, d\underline{x} .$$

We need to evaluate integrals of the following type:

$$\int_{[R]} \cdots \int (\sum_1^n x_i^2)^k \quad d\underline{x} .$$

If we expand the integrand according to the multinomial theorem, we get

$$\sum_{i_1+\cdots+i_n=k} \frac{k!}{i_1! \cdots i_n!} \int_{[R]} \cdots \int \prod_{j=1}^{n} x_j^{2i_j} \, dx_j .$$

We now make use of the Dirichlet integral stated earlier, where

$c_j = (\frac{2t}{a_j})^{1/2}$, $p_j = 2$, $\ell_j = (2i_j+1)$, getting

$$\int\cdots\int_{[R]} (\sum_1^n X_i^2)\, d\underline{x} = \frac{(2t)^{k+\frac{n}{2}} k!}{(a_1\ldots a_n)^{1/2}\Gamma(\frac{n}{2}+k+1)} \cdot$$

$$\sum_{i_1+\ldots+i_n=k} \frac{\Gamma(i_1+\frac{1}{2})\ldots\Gamma(i_n+\frac{1}{2})}{i_1!\ldots i_n! a_1^{i_1}\ldots a_n^{i_n}}.$$

The problem now is to evaluate this last expression. Recalling that if X_i is $N(0, 1)$, then $E(X_i^2)^k = \frac{2^k\Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})}$, we find that

$$E(Q_n^*)^k = E\left[\frac{1}{2}\sum_1^n a_i^{-1}X_i^2\right]^k$$

$$= 2^{-k}E \sum_{i_1+\ldots+i_n=k} \frac{k!}{i_1!\ldots i_n!} \frac{X_1^{2i_1}\ldots X_n^{2i_n}}{a_1^{i_1}\ldots a_n^{i_n}}$$

$$= 2^{-k} k! \sum_{i_1+\ldots+i_n=k} \frac{E X_1^{2i_1} \ldots E X_n^{2i_n}}{i_1! \ldots i_n! \; a_1^{i_1} \ldots a_n^{i_n}}$$

$$= \pi^{-\frac{n}{2}} k! \sum_{i_1+\ldots+i_n=k} \frac{\Gamma(i_1+\frac{1}{2}) \ldots \Gamma(i_n+\frac{1}{2})}{i_1! \ldots i_n! \; a_1^{i_1} \ldots a_n^{i_n}} .$$

So that

$$\underset{[R]}{\int \cdots \int} \left(\sum_1^n x_i^2\right)^k d\underline{x} = \frac{(2t)^{k+\frac{n}{2}} \pi^{\frac{n}{2}}}{(a_1 \ldots a_n)^{1/2} \, \Gamma(\frac{n}{2}+k+1)} \quad E(Q_n^*)^k ,$$

and

$$F_n(t) = \frac{t^{n/2}}{(a_1 \ldots a_n)^{1/2}} \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} \frac{E(Q_n^*)^k}{\Gamma(\frac{n}{2}+k+1)} .$$

This proves part (a) of the theorem.

To show absolute convergence we note that if

$$a_1 \geq a_2 \geq \ldots \geq a_n > 0, \text{ then}$$

$$Q_n^* = \frac{1}{2}\left(\frac{X_1^2}{a_1} + \ldots + \frac{X_n^2}{a_n}\right) \leq \frac{1}{2a_n} \sum_1^n X_i^2, \text{ and}$$

$$E(Q_n^*)^k \leq \frac{1}{a_n^k} \; E(\tfrac{1}{2}\sum_1^n X_i^2)^k = \frac{\Gamma(\frac{n}{2}+k)}{a_n^k\,\Gamma(\frac{n}{2})}, \text{ so that}$$

$$F_n(t) \leq \frac{t^{n/2}}{(a_1\ldots a_n)^{1/2}} \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{\Gamma(\frac{n}{2}+k)}{\Gamma(\frac{n}{2}+k+1)\,a_n^k\,\Gamma(\frac{n}{2})}$$

$$\leq \frac{t^{n/2}}{(a_1\ldots a_n)^{1/2}} \frac{e^{t/an}}{\Gamma(\frac{n}{2}+1)} < \infty, \text{ for finite } t.$$

This proves part (b) of the theorem.

The bounds we obtain are based on the fact that if r and s are any two positive integers ≥ 1, then for $z > 0$, $\sum_{k=0}^{2s-2} \frac{(-z)^k}{k!} > e^{-z} > \sum_{k=0}^{2r-1} \frac{(-z)^k}{k!}$.

This proves part (c) of the theorem.

In the case where some of the latent roots are zero, i.e. when the form is positive semi-definite of rank r, say, we need only replace n by r in the theorem and in the proof.

<u>Remarks.</u>

(i) The moments of Q_n^*, $E(Q_n^*)^k$, are easy to obtain from the cumulants of Q_n^*. The r-th cumulant of Q_n^* is $k_r(Q_n^*) = \frac{(r-1)!}{2} \sum_1^n a_i^{-r}$.

From Kendall [11], we have expressions for the first ten moments in terms of the cumulants.

(ii) Let S_r be the sum of the first r+1 terms of the series for $F_n(t)$. Then $S_0, S_2, S_4, \ldots$ is a sequence of upper bounds and $S_1, S_3, S_5, \ldots$ is a sequence of lower bounds to $F_n(t)$. If E_{2k+2} is the error committed by stopping with S_{2k+1}, then

$$E_{2k+2} \leq \text{l.u.b. } S_{2r} - \text{g.l.b. } S_{2r+1}, \qquad r=0,1,\ldots,k;\ k=0,1,\ldots,$$

$$E_{2k+1} \leq \text{l.u.b. } S_{2r} - \text{g.l.b. } S_{2r-1}, \qquad r=1,2,\ldots,k;\ k=1,2,\ldots \ .$$

The values of l.u.b. S_{2r} and g.l.b. S_{2r+1} depend on the values of the latent roots. We note that $E_{2k+2} \leq S_{2k} - S_{2k+1}$ = the last term included, and $E_{2k+1} \leq S_{2k} - S_{2k-1}$ = the last term included. Hence, the error is less than the last term included and it is positive if we take an odd number of terms and negative if we take an even number of terms.

(iii) The above theorem seems to be in several ways an improvement over the method given in Robbins [16].

1.3 An application: The distribution of a sum of squares in dependent variates.

Suppose that $X_1, \ldots, X_n$ have a joint multivariate normal distribution with zero means and covariance matrix equal to A^{-1}.

Then the problem is, what is the distribution of $\frac{1}{2}\sum_{1}^{n} X_i^2$? Now

$$\Pr\left[\tfrac{1}{2} X'X \leq t\right] = \frac{|A|^{1/2}}{(2\pi)^{n/2}} \int_{[\frac{1}{2} X'X \leq t]} \cdots \int \operatorname{Exp} - \tfrac{1}{2} X'A\,X \, d\underline{x}.$$

Make an orthogonal transformation, $X = L\,Y$, say where $L\,L' = L'L = I$.

Then $X'X = Y'Y$ and $X'A\,X = Y'L'A\,L\,Y = \sum_{1}^{n} a_i Y_i^2$, where $a_1, \ldots, a_n$ are

the latent roots of the matrix A. Now make the transformation

$z_i^2 = a_i Y_i^2$, getting

$$\Pr\left[\tfrac{1}{2} X'X \leq t\right] = (2\pi)^{-\frac{n}{2}} \int_{[\frac{1}{2}\sum a_i^{-1} z_i^2 \leq t]} \cdots \int \operatorname{Exp} - \tfrac{1}{2} z'z \, d\underline{z}$$

$$= \Pr\left[Q_n^* \leq t\right],$$ and we can make use of

theorem 1.1.

<u>Remark</u>.

Combining the results of theorem 1.1 and the above application it is easy to show that we could find the distribution of a definite q.f., $X'A\,X$, in $X_1, \ldots, X_n$ where $X_1, \ldots, X_n$ have a multivariate normal distribution with covariance matrix B^{-1}, and this distribution involves as parameters the characteristic roots of AB^{-1}.

We shall now state, without proof, an obvious corollary to theorem 1.1, obtained by letting the first m_1 latent roots be a_1, the next m_2 latent roots be a_2, etc.

Corollary 1.1

Let $S_r = \frac{1}{2}(a_1 \chi^2_{m_1} + \ldots + a_r \chi^2_{m_r})$, where the $\chi^2_{m_i}$ are independent r.v.'s having a central chi-square distribution with m_i degrees of freedom. Let $S^*_r = \frac{1}{2}(a_1^{-1} \chi^2_{m_1} + \ldots + a_r^{-1} \chi^2_{m_r})$, $M = \sum_1^r m_i$, $a_i > 0$, $G_r(t) = \Pr(S_r \leq t)$, then

$$G_r(t) = \frac{t^{M/2}}{(a_1^{m_1} \ldots a_r^{m_r})^{1/2}} \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} \frac{E(S^*_r)^k}{\Gamma(\frac{M}{2}+k+1)},$$

where $E(S^*_r)^k$ is the k-th moment of S^*_r. The series is absolutely convergent, and for any two positive integers s and j and every $t > 0$,

$$\sum_{k=0}^{2s-2} d_k > G_r(t) > \sum_{k=0}^{2j-1} d_k, \text{ where}$$

$$d_k = \frac{t^{M/2}}{(a_1^{m_1}\cdots a_r^{m_r})^{1/2}} \frac{(-t)^k}{k!} \frac{E(S_r^*)^k}{\Gamma(\frac{M}{2}+k+1)} .$$

Remarks.

(i) The moments of S_r^* are easy to obtain from the cumulants. The j-th cumulant of S_r^* is $k_j(S_r^*) = \frac{(j-1)!}{2} \sum_{i=1}^{r} m_i a_i^{-j}$. We can find the first ten moments in terms of the cumulants from Kendall [11].

(ii) The above corollary gives a method which seems to be an improvement over the one given in Robbins and Pitman [17].

1.4 Hotelling's method of Laguerre polynomials.

In this section we shall give a very brief account of a method suggested by Hotelling [10]. Let $Q_n = \frac{1}{2}(a_1X_1^2+\ldots+a_nX_n^2)$, where the X_i are independent $N(0, 1)$ variates and where $a_i > 0$, $i=1,\ldots,n$. Let $g(q)$ be the p.d.f. of Q_n and let

$$f(x) = \frac{e^{-x}x^{m-1}}{\Gamma(m)} \quad , \quad x > 0 \quad , \quad \text{where } m = \frac{n}{2}.$$

Then the suggested expansion for $g(q)$ is the Laguerre series

$$g(q) = f(q) \sum_{r=0}^{\infty} b_r L_r^{(m-1)}(q),$$

where $L_r^{(m-1)}(q)$, $r=0, 1, \ldots,$ is the sequence of Laguerre polynomials satisfying the relation

$$\int_0^{\infty} f(x) L_i^{(m-1)}(x) L_j^{(m-1)}(x)\, dx = \binom{i+m-1}{i} \delta_{ij}$$

$$i, j = 0, 1, \ldots, \text{ for } m > 0.$$

The Laguerre polynomial $L_s^{(m-1)}(q)$ has the following explicit representation:

$$L_s^{(m-1)}(q) = \sum_{v=0}^{s} \binom{s+m-1}{s-v} \frac{(-q)^v}{v!}, \qquad s=0,1,\ldots \quad .$$

See Szego [18]. It follows from the orthogonality condition that

$$b_s = \frac{s!\, \Gamma(m)}{\Gamma(s+m)} \int_0^{\infty} g(q)\, L_s^{(m-1)}(q)\, dq,$$

and so b_s is a linear function of the moments of Q_n. The series

$$g(q) = f(q)\, (1+b_1 L_1(q)+b_2 L_2(q)+\ldots)$$

converges uniformly over the whole real axis and so we can integrate term by term.

Remarks.

(i) The main drawback in using the above expansion is that no convenient bound is known for the error committed by stopping after a certain number of terms.

(ii) Hotelling suggests that the above series be used to find the p.d.f. of the ratio of a definite quadratic form to a sum of squares and the p.d.f. of an indefinite form by convolution, since an indefinite form is the difference of two definite forms.

CHAPTER II

THE DISTRIBUTION OF A DEFINITE QUADRATIC FORM IN INDEPENDENT NON-CENTRAL NORMAL VARIATES

2.1 The problem.

Suppose we have a q. f. $Q_n = \frac{1}{2} Y'AY$ in $Y_1, \ldots, Y_n$ where the Y_i are independent $N(\xi_i, 1)$ variates, and where $Y' = (Y_1, \ldots, Y_n)$. Let

$$G_n(t; \xi_1, \ldots, \xi_n) = G_n(t; \underline{\xi}) = Pr(Q_n \leq t)$$

then the problem is to find $G_n(t; \underline{\xi})$. Let us make an orthogonal transformation $Y = \Gamma X$, say, where $\Gamma'\Gamma = \Gamma\Gamma' = I$, such that

$$Y'AY = X'\Gamma'A\Gamma X = X'D_a X = \sum_1^n a_i X_i^2 \quad ,$$

where D_a is a diagonal matrix having the elements $a_1, \ldots, a_n$ in the main diagonal, where $a_1, \ldots, a_n$ are the latent roots of the matrix A. If $EY = \xi$, then $EX = \Gamma'\xi = \mu$ (say), where $\xi' = (\xi_1, \ldots, \xi_n)$ and $\mu' = (\mu_1, \ldots, \mu_n)$. Hence, under such a transformation $X_1, \ldots, X_n$ remain independent with the same variances as $Y_1, \ldots, Y_n$, but the

mean of X_i is now $\mu_i = (\gamma_{1i}\xi_1 + \ldots + \gamma_{ni}\xi_n)$, where $(\gamma_{1i}, \ldots, \gamma_{ni})$ are the elements of the i-th row of Γ'.

So the problem is now to find the $\Pr \left[\frac{1}{2} \sum_1^n a_i X_i^2 \leq t \right]$, where the X_i are independent $N(\mu_i, 1)$.

2.2 The solution.

Theorem 2.1 Let $Q_n = \frac{1}{2} \sum_1^n a_i X_i^2$, where the X_i are independent $N(\mu_i, 1)$ and where $a_1 \geq a_2 \geq \ldots \geq a_n > 0$. Let $Q_n^{**} = \frac{1}{2} \sum_{k=1}^n a_n^{-1} Y_k^2$, where the Y_k are independent $N(i\mu_k, 1)$, $i = \sqrt{-1}$. Then

(a) $$G_n(t; \underline{\mu}) = \frac{e^{-\lambda} t^{n/2}}{(a_1 \ldots a_n)^{1/2}} \sum_{s=0}^{\infty} \frac{(-t)^s}{s!} \frac{c_s}{\Gamma(\frac{n}{2} + s + 1)},$$

where $c_s = E(Q_n^{**})^s$, and $\lambda = \frac{1}{2} \sum_{i=1}^n \mu_i^2$.

(b) The series is absolutely convergent and therefore it is convergent.

(c) For any two positive integers r and k and every $t > 0$,

$$\sum_{s=0}^{2r} d_s > G_n(t;\underline{\mu}) > \sum_{s=0}^{2k-1} d_s, \text{ where}$$

$$d_s = \frac{e^{-\lambda}\, t^{n/2}}{(a_1 \ldots a_n)^{1/2}} \quad \frac{(-t)^s}{s!} \quad \frac{c_s}{\Gamma(\frac{n}{2} + s + 1)}$$

Proof. We know that if the X_i are independent $N(\mu_i, 1)$ and if

$\lambda = \frac{1}{2} \sum_1^n \mu_i^2$, and $Y = \frac{1}{2} \sum_1^n X_i^2$, then the p.d.f. of Y is

$$e^{-\lambda}\, e^{-y}\, y^{\frac{n}{2}-1} \sum_{m=0}^{\infty} \frac{(\lambda y)^m}{m!\, \Gamma(m + \frac{n}{2})} \quad .$$

Let $[R] = [\sum_1^n a_i y_i \leq t]$, $d\underline{y} = dy_1 \ldots dy_n$, and $\lambda_j = \frac{1}{2}\mu_j^2$, then

$$G_n(t;\mu) = \int_{[R]}\!\!\cdots\!\int \prod_{j=1}^{n} e^{-\lambda_j}\, e^{-y_j}\, y_j^{-\frac{1}{2}} \sum_{i_j=0}^{\infty} \frac{(\lambda_j y_j)^{i_j}}{i_j!\, \Gamma(i_j + \frac{1}{2})}\, dy_j$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \sum_{i_1+\cdots+i_n=k} \frac{\lambda_1^{i_1} \cdots \lambda_n^{i_n}}{i_1! \cdots i_n! \, \Gamma(i_1+\frac{1}{2}) \cdots \Gamma(i_n+\frac{1}{2})} \cdot I(t;\underline{a};n) ,$$

where

$$I(t;\underline{a};n) = \int \cdots \int_{[R]} \operatorname{Exp}\left[- \sum_1^n y_i \right] \prod_{j=1}^{n} y_j^{i_j - \frac{1}{2}} \, dy_j \quad .$$

Expanding the exponential and making use of the Dirichlet integral stated in (1.2), we obtain:

$$G_n(t;\underline{\mu}) = e^{-\lambda} \sum_{k=0}^{\infty} \sum_{i_1+\cdots+i_n=k} \frac{\lambda_1^{i_1} \cdots \lambda_n^{i_n}}{i_1! \cdots i_n! \; \Gamma(i_1+\frac{1}{2}) \cdots \Gamma(i_n+\frac{1}{2})} \cdot$$

$$\sum_{r=0}^{\infty} \sum_{j_1+\cdots+j_n=r} \frac{(-1)^r t^{r+k+\frac{n}{2}}}{j_1! \cdots j_n!} \; \frac{\Gamma(i_1+j_1+\frac{1}{2}) \cdots \Gamma(i_n+j_n+\frac{1}{2})}{a_1^{i_1+j_1+\frac{1}{2}} \cdots a_n^{i_n+j_n+\frac{1}{2}} \; \Gamma(k+r+\frac{n}{2}+1)} \quad .$$

This can be rewritten as

$$G_n(t;\,\underline{\mu}) = \frac{e^{-\lambda}\,t^{n/2}}{(a_1\ldots a_n)^{1/2}} \sum_{s=0}^{\infty} c_s \frac{(-t)^s}{s!\,\Gamma(\frac{n}{2}+s+1)}\;, \text{ where}$$

$$c_s = \sum_{k=0}^{s} \sum_{i_1+\ldots+i_n=k} \sum_{j_1+\ldots+j_n=s} d,$$

where $d = \dfrac{s!\,\lambda_1^{i_1}\ldots\lambda_n^{i_n}\,\Gamma(j_1+\frac{1}{2})\ldots\Gamma(j_n+\frac{1}{2})(-1)^k}{a_1^{j_1}\ldots a_n^{j_n}\,i_1!\ldots i_n!\,(j_1-i_1)!\ldots(j_n-i_n)!\,\Gamma(i_1+\frac{1}{2})\ldots\Gamma(i_n+\frac{1}{2})}$.

The problem now is to evaluate c_s.

Let $Q_n^{**} = \frac{1}{2}\sum_{k=1}^{n} a_k^{-1}Y_k^2$, where the Y_k are independent $N(i\mu_k,\,1)$, $i=\sqrt{-1}$,

then

$$EY^{2r} = (2\pi)^{-1/2} \int_{-\infty}^{\infty} Y^{2r}\,\mathrm{Exp}\,-\tfrac{1}{2}(Y-i\mu)^2\,dY$$

$$= \pi^{-1/2} \sum_{j=0}^{r} \binom{2r}{2j} (-\mu^2)^j \, 2^{r-j} \, \Gamma(r-j+\tfrac{1}{2})$$

$$= \sum_{j=0}^{r} \binom{r}{j} \frac{\Gamma(r+\frac{1}{2})}{\Gamma(j+\frac{1}{2})} (-\mu^2)^j \quad 2^{r-j} \qquad , \quad \text{and}$$

$$E(Q_n^{**})^s = 2^{-s} \sum_{j_1+\ldots+j_n=s} \frac{s!}{j_1! \ldots j_n!} \frac{E\, Y_1^{2j_1} \ldots Y_n^{2j_n}}{a_1^{j_1} \ldots a_n^{j_n}}$$

$$= 2^{-s} \sum_{j_1+\ldots+j_n=s} \frac{s!\, 2^{s-k}}{j_1! \ldots j_n! a_1^{j_1} \ldots a_n^{j_n}} \sum_{i_1=0}^{j_1} \ldots \sum_{i_n=0}^{j_n} h \, ,$$

$$\text{where } h = \binom{j_1}{i_1} \ldots \binom{j_n}{i_n} \frac{\Gamma(j_1+\frac{1}{2})\ldots\Gamma(j_n+\frac{1}{2})(-\mu_1^2)^{i_1}\ldots(-\mu_n^2)^{i_n}}{\Gamma(i_1+\frac{1}{2}) \ldots \Gamma(i_2+\frac{1}{2})}$$

$$= \sum_{k=0}^{s} \sum_{i_1+\ldots+i_n=k} \sum_{j_1+\ldots+j_n=s} d$$

where $d = \dfrac{s!\,\lambda_1^{i_1}\ldots\lambda_n^{i_n}(-1)^k\,\Gamma(j_1+\frac{1}{2})\ldots\Gamma(j_n+\frac{1}{2})}{i_1!\ldots i_n!(j_1-i_1)!\ldots(j_n-i_n)!\,\Gamma(i_1+\frac{1}{2})\ldots\Gamma(i_n+\frac{1}{2})a_1^{j_1}\ldots a_n^{j_n}}$.

But this is just c_s. Hence,

$$G_n(t;\underline{\mu}) = \frac{e^{-\lambda}\,t^{n/2}}{(a_1\ldots a_n)^{1/2}} \sum_{s=0}^{\infty} \frac{(-t)^s}{s!} \frac{E(Q_n^{**})^s}{\Gamma(\frac{n}{2}+s+1)}$$

This proves (a) of the theorem.

If $k_r(Q_n)$ is the r-th cumulant of Q_n, it is not difficult to show that

$$k_r(Q_n) = \frac{(r-1)!}{2} \sum_{j=1}^{n} a_j^r(1+r\mu_j^2) .$$

Hence, to find the r-th cumulant of Q_n^{**} we must replace μ_j by $i\mu_j$ and a_j by a_j^{-1}, getting, $k_r(Q_n^{**}) = \frac{(r-1)!}{2} \sum_{j=1}^{n} a_j^{-r}(1-r\mu_j^2)$. Therefore, c_s is the s-th moment of a r.v. whose s-th cumulant is $k_s(Q_n^{**})$.

If $Q_n^* = \frac{1}{2} \sum_1^n a_i^{-1} X_i^2$, where the X_i are independent $N(0, 1)$,

then $\quad k_r(Q_n^*) = \frac{(r-1)!}{2} \sum_{j=1} a_j^{-r}$. Therefore, $k_r(Q_n^{**}) \leq k_r(Q_n^*)$,

and it follows that $E(Q_n^{**})^s \leq E(Q_n^*)^s \leq \frac{\Gamma(\frac{n}{2} + s)}{a_n^s \Gamma(\frac{n}{2})}$. Consequently,

$$G_n(t; \underline{\mu}) \leq \frac{e^{-\lambda} t^{n/2}}{(a_1 \ldots a_n)^{1/2}} \sum_{s=0}^{\infty} \frac{t^s}{s!} \frac{\Gamma(\frac{n}{2} + s)}{a_n^s \Gamma(\frac{n}{2} + s + 1)}$$

$$\leq \frac{e^{-\lambda} t^{n/2} e^{t/a_n}}{(a_1 \ldots a_n)^{1/2} \Gamma(\frac{n}{2} + 1)} < \infty \text{ for finite } t \text{ .}$$

This proves (b) of the theorem. The bounds stated in theorem 2.1 are based on the fact that if r and s are any two positive integers ≥ 1, then for $z > 0$,

$$\sum_{k=0}^{2s} \frac{(-z)^k}{k!} > e^{-z} > \sum_{k=0}^{2r-1} \frac{(-z)^k}{k!} \quad .$$

This proves (c) of the theorem.

Theorem 2.2. Under the conditions of theorem 2.1,

$$G_n(t;\ \underline{\mu}) \geq e^{-\lambda}\, G_n(t;\ \underline{0}) \quad ,$$

where $G_n(t;\ \underline{0})$ equals the $F_n(t)$ of theorem 1.1.

Proof.

$$G_n(t;\underline{\mu}) = \int \cdots \int_{[R]} \prod_{i=1}^{n} e^{-\lambda_i} \quad e^{-y_i} \quad y_i^{-1/2} \sum_{j_i=0}^{\infty} \frac{(\lambda_i y_i)^{j_i}}{j_i!\,\Gamma(j_i+\frac{1}{2})} \; dy_i .$$

Now,

$$\int \cdots \int_{[R]} \prod_{i=1}^{n} \frac{e^{-y_i}\, y_i^{\frac{2j_i+1}{2}-1}}{\Gamma(\frac{2j_i+1}{2})} \, dy_i = \Pr\left[\tfrac{1}{2}(a_1 \chi^2_{2j_1+1} + \cdots + a_n \chi^2_{2j_n+1}) \leq t\right],$$

where $\chi^2_{2j_m+1}$ is a r.v. having a central chi-square distribution

with $2j_m+1$ degrees of freedom. Hence,

$$G_n(t;\underline{\mu}) = e^{-\lambda} \sum_{k=0}^{\infty} \sum_{j_1+\cdots+j_n=k} \frac{\lambda_1^{j_1} \cdots \lambda_n^{j_n}}{j_1! \cdots j_n!} \quad .$$

$$\Pr\left[\tfrac{1}{2}(a_1 \chi^2_{2j_1+1} + \cdots + a_n \chi^2_{2j_n+1}) \leq t\right],$$

and so,

$$G_n(t; \underline{\mu}) \geq e^{-\lambda}\, G_n(t; \underline{0}) \quad .$$

We note that equality is attained for any given t, by putting $\lambda = 0$ (in which case $\underline{\mu} = \underline{0}$) .

Theorem 2.3. Under the conditions of theorem 2.1,

$$G_n(t; \underline{\mu}) \leq G_n(t; \underline{0}) \quad .$$

Proof.

Using a special case of a more general unpublished result due to Sigeity Moriguti, we find that

$$\int_{[ax^2 \leq t]} \exp - \tfrac{1}{2}(x-\mu)^2 \, dx \leq \int_{[ax^2 \leq t]} \exp - \tfrac{1}{2} x^2 \, dx \quad .$$

We can generalize the above inequality to the case n = 2 as follows:

$$\int\int_{[a_1x_1^2+a_2x_2^2 \leq t]} \exp - \tfrac{1}{2} [(x_1-\mu_1)^2 + (x_2-\mu_2)^2] \; dx_1 \, dx_2$$

$$= \int_{[a_2x_2^2 \leq t]} \int_{[a_1x_1^2 \leq t-a_2x_2^2]} \exp - \tfrac{1}{2}(x_1-\mu_1)^2 dx_1 \;\; \exp - \tfrac{1}{2}(x_2-\mu_2)^2 \, dx_2$$

$$\leq \int_{[a_2x_2^2 \leq t]} \int_{[a_1x_1^2 \leq t - a_2x_2^2]} \exp - \tfrac{1}{2}x_1^2 dx_1 \quad \exp - \tfrac{1}{2}(x_2-\mu_2)^2 dx_2$$

$$\leq \int_{[a_2x_2^2 \leq t]} \int_{[a_1x_1^2 \leq t - a_2x_2^2]} \exp - \tfrac{1}{2}x_1^2 dx_1 \quad \exp - \tfrac{1}{2}x_2^2 dx_2$$

$$= \int\int_{[a_1x_1^2 + a_2x_2^2 \leq t]} \exp - \tfrac{1}{2}(x_1^2+x_2^2)\, dx_1\, dx_2 \quad .$$

Generalizing the method used for the case $n=2$, it is not difficult to show that the above inequality could be established for $n=3,4,\ldots$. This completes the proof. Combining the results of theorems 2.2 and 2.3 we have the very useful inequality:

$$G_n(t;\underline{0}) \geq G_n(t;\underline{\mu}) \geq e^{-\lambda}\, G_n(t;\underline{0}).$$

The proof of theorem 2.3 was suggested to the writer by Professor S. N. Roy.

Remarks.

(i) The introduction of an imaginary mean in theorem 2.1 is merely a mathematical convenience; we could have omitted this formulation altogether and merely stated that c_s is the s-th

moment of a r.v. whose s-th cumulant is $k_s(Q_n^{**})$.

(ii) The moments c_s are easy to compute when we know the cumulants. Kendall [11], gives the first ten moments in terms of the cumulants.

(iii) It can be shown that $G_n(t;\mu)$ is invariant under an orthogonal transformation $Y=\Gamma X$, say, such that $\Gamma' A \Gamma = D_a$ and $\Gamma'\Gamma = \Gamma\Gamma' = I$.

2.3 An application: The power function of the chi-square statistic.

Suppose that the observations from a random experiment can fall into any one of k cells and that the expected number of observations in the i-th cell under H_0 and H_1 is m_i^0 and m_i, respectively.

That is: H_0: $m_1^0,\ldots,m_k^0$, and H_1: $m_1,\ldots,m_k$. Let

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(n_i-m_i^0)^2}{m_i^0}, \quad \text{and}$$

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i-m_i)^2}{m_i}, \quad \text{where}$$

n_i is the observed number of observations in the i-th cell. If $1-\beta$ is the power of the test, then

$$\beta = \Pr[\, X_0^2 \le t \mid m_1,\ldots,m_k \,]$$

$$= (2\pi)^{-k/2} \int\cdots\int_{[X_0^2 \le t]} \exp[-\tfrac{1}{2} X^2] \prod_{i=1}^{k} d\left(\frac{n_i - m_i}{\sqrt{m_i}}\right)$$

approximately, for sufficiently large n_i. Putting $x_i = \dfrac{n_i - m_i}{\sqrt{m_i}}$,

$a_i/2 = m_i/m_i^0$, $\mu_i = \dfrac{m_i - m_i^0}{\sqrt{m_i}}$, and letting $x_i + \mu_i = y_i$, we get

$$\Pr[X_0^2 \le t \mid m_1,\ldots,m_k] = (2\pi)^{-(k/2)} \int\cdots\int_{[\frac{1}{2}\sum_1^k a_i y_i^2 \le t]} \exp[-\tfrac{1}{2}\sum_1^k (y_i - \mu_i)^2] \prod_1^k dy_i$$

$$= G_k(t;\underline{\mu}) \quad .$$

Hence, the application of theorem 2.1 will give us the power function of the chi-square statistic.

CHAPTER III

THE DISTRIBUTION OF AN INDEFINITE QUADRATIC FORM IN INDEPENDENT CENTRAL NORMAL VARIATES

3.1 Special case 1.

As a preliminary step to the study of the distribution of an indefinite q.f., this section is concerned with the distribution of the difference between two independent chi-squares having different numbers of degrees of freedom. If the degrees of freedom are the same, the distribution becomes the same as that of the sample covariance in sampling from a normal population. We shall study some of the properties of this distribution in order to anticipate the behavior of the distribution of a more general indefinite q.f. Others [15], [4], [2] have considered this problem from a slightly different standpoint. The main results of this section are:

(i) Recurrence relations 3.1.4, 3.1.5, 3.1.6, 3.1.7, 3.1.20

(ii) Inequalities 3.1.9, 3.1.21

(iii) Further properties 3.1.12, 3.1.13, 3.1.15, 3.1.16

(iv) An application of a result due to Berry [3], 3.1.19 .

If $T_{n,m} = X_n - Y_m$, where X_n and Y_m are independently distributed with p.d.f. $h_n(x)$ and $h_m(x)$ respectively, where

$$h_n(x) = \frac{1}{\Gamma(\frac{n}{2})} e^{-x} x^{\frac{n}{2}-1}, \quad x > 0,$$

then if the p.d.f. of $T_{n,m}$ is $g_{n,m}(t)$,

$$g_{n,m}(t) = \int_0^{\infty} h_n(x+t)h_m(x)\,dx, \qquad t > 0,$$

$$= \int_0^{\infty} h_n(x)\,h_m(x-t)\,dx, \qquad t < 0.$$

We see that given the p.d.f. for $t > 0$, to get it for $t < 0$, replace t by $-t$ and interchange n and m. Therefore, we shall consider only the case $t > 0$, for definiteness.

Hence,

3.1.1 $$g_{n,m}(t) = \frac{e^{-t}}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} \int_0^{\infty} e^{-2x}(x+t)^{\frac{n}{2}-1} x^{\frac{m}{2}-1}\,dx, \quad t > 0.$$

The moment generating function of $T_{n,m}$ is $M(\theta) = Ee^{t\theta} = (1-\theta)^{-\frac{n}{2}}(1+\theta)^{-\frac{m}{2}}$. From $M(\theta)$ we see that

(i) If $n,m \longrightarrow \infty$ so that $\frac{n}{m} \longrightarrow 1$, then $T_{n,m}$ is asymptotically normal $(\frac{n-m}{2}, \sqrt{\frac{n+m}{2}})$.

(ii) If $n \longrightarrow \infty$, but m remains finite, then $T_{n,m}$ is

asymptotically normal $(\frac{n-m}{2}), \sqrt{n})$.

(iii) If $m \longrightarrow \infty$, but n remains finite, then $T_{n,m}$ is

asymptotically normal $(\frac{n-m}{2}, \sqrt{m})$.

In 3.1.1 let $x/t = y$, getting

$$3.1.2 \qquad g_{n,m}(t) = \frac{e^{-t} t^{\frac{n+m}{2}-1}}{\Gamma(\frac{n}{2})\,\Gamma(\frac{m}{2})} \int_0^{\infty} e^{-2yt}(1+y)^{\frac{n}{2}-1} y^{\frac{m}{2}-1}\,dy.$$

Let $n/2 = p$, $m/2 = q$, and let

$$3.1.3 \qquad I(p,q) = \int_0^{\infty} e^{-2yt}(1+y)^p\, y^q\,dy.$$

Integrate 3.1.3 by parts three separate times as follows:

(i) $u = (1+y)^p\, y^q$, $\qquad dv = e^{-2yt}\,dy$,

(ii) $u = (1+y)^p\, e^{-2yt}$, $\qquad dv = y^q\,dy$,

(iii) $u = (1+y)^p\, y^{q-1}$, $\qquad dv = e^{-2yt} y\,dy$, getting

$$(i)' \qquad I(p,q) = (\tfrac{1}{2t})\,[\,q\, I(p,q-1) + p\, I(p-1,q)\,],$$

$$m > 0,\ q > 0,$$

$$(ii)' \quad I(p,q) = -\left(\frac{p}{q+1}\right) I(p-1,q+1) + \left(\frac{2t}{q+1}\right) I(p,q+1),$$

$$m > -2, \; q+1 > 0,$$

$$(iii)' \quad I(p,q) = \left(\frac{1}{4t^2}\right) \left[(q-1)\, I(p,q-2) + p\, I(p-1,q-1) \right.$$

$$\left. + 2t(q-1)\, I(p,q-1) + 2tp\, I(p-1,q) \right],$$

$m > 2$, $q-1 > 0$, respectively.

The restrictions placed on m and q are to prevent the integrated part from becoming infinite. Equate (i)' to (iii)' getting

$$(iv) \quad I(p+1,q) = \left(\frac{q}{2t}\right) I(p+1,q-1) + \left(\frac{p+1}{2t}\right) I(p,q).$$

In (i)' replace p by p+1 and q by q+1 and then substitute from (ii)' and (iv) into (i)' getting

$$(v) \quad 4t^2\, I(p+1,q+1) = 2(p+1)(q+1)\, I(p,q) + q(q+1)\, I(p+1,q-1)$$

$$+ p(p+1)\, I(p-1,q+1).$$

Now $q(q+1)\, I(p+1,q-1) + p(p+1)\, I(p-1,q+1) = q(q+1)\, I(p-1,q-1)$ $+ \left[p(p+1)+q(q+1) \right] I(p-1,q+1) + 2q(q+1)\, I(p-1,q)$, and $I(p-1,q+1)$ $+I(p-1,q) = I(p,q)$. Substitute into (v) getting

$$4t^2 \, I(p+1,q+1) = [2(p+1)(q+1)+p(p+1)+q(q+1)] \, I(p,q)$$

$$+ \, q(q+1) \, I(p-1,q-1) + [q(q+1)-p(p+1)] \, I(p-1,q),$$

so that finally,

3.1.4 $$g_{n+4,m+4}(t) = \left[\frac{n(n+2)+m(m+2)+2(n+2)(m+2)}{4(n+2)(m+2)}\right] g_{n+2,\, m+2}(t)$$

$$+ \frac{t^2}{n(n+2)} \; g_{n,m}(t) + \left[\frac{m(m+2)-n(n+2)}{2n(n+2)(m+2)}\right] t \; g_{n,m+2}(t); \quad t > 0.$$

The same relation holds for $t < 0$ if we replace t by $-t$ and interchange n and m. Note that if $n=m$, the last term vanishes and the relation reduces to

$$g_{n+4,n+4}(t) = \left(\frac{n+1}{n+2}\right) g_{n+2,n+2}(t) + \frac{t^2}{n(n+2)} \, g_{n,m}(t).$$

From the first integration by parts we get the simple relation

3.1.5 $$g_{n+2,m+2}(t) = \tfrac{1}{2}[g_{n+2,m}(t) + g_{n,m+2}(t)], \text{ for all } t.$$

We now make use of the p.d.f.'s to obtain the c.d.f.'s. For $x > 0$,

$$\Pr[T_{n,m} > x] = \int_x^{\infty} g_{n,m}(t) \, dt$$

$$= \frac{1}{\Gamma(\frac{n}{2})\,\Gamma(\frac{m}{2})} \int_x^{\infty} e^{-t} \int_0^{\infty} e^{-2y}(y+t)^{\frac{n}{2}-1} y^{\frac{m}{2}-1} dy\, dt.$$

Integrate by parts where $dv = e^{-t}dt$, and $u = \int_0^{\infty} e^{-2y}(y+t)^{\frac{n}{2}-1} y^{\frac{m}{2}-1} dy$,

getting

$$Pr[T_{n,m} > x] = g_{n,m}(x) + Pr[T_{n-2,m} > x].$$

Hence, if n is even,

3.1.6 $$Pr[T_{n,m} > x] = \sum_{j=0}^{(n-2)/2} g_{n-2j,m}(x), \text{ where}$$

$$g_{2,m}(x) = Pr[T_{2,m} > x] = \frac{e^{-x}}{2^{m/2}} .$$

If n is odd

3.1.7 $$Pr[T_{n,m} > x] = \sum_{j=0}^{(n-3)/2} g_{n-2j,m}(x) + Pr[T_{1,m} > x] .$$

From 3.1.6 and 3.1.7 it is seen that if we have a table of $g_{n,m}(t)$ and if we know $Pr[T_{1,m} > x]$ for all m, we can find $Pr[T_{n,m} > x]$

for all n and m.

Our first objective, then, is to be able to find $g_{n,m}(t)$ for any n and m. If we consider a rectangular table of $g_{n,m}(t)$ having n columns and m rows, we will find that given $g_{n,m}(t)$ for n,m=1,2,3,4,5 and $g_{2,m}(t)$ for m=1,2,..., we can complete the table using 3.1.4 and 3.1.5.

First of all we can fill in the second column using the fact that

$$g_{2,m}(t) = \frac{e^{-t}}{2^{m/2}} , \qquad m=1,2,..., \qquad \text{and}$$

$$g_{4,m}(t) = \frac{e^{-t}}{2^{m/2}} \left(t + \frac{m}{4}\right) , \qquad m=1,2,3,4,5 \quad .$$

If we let n=m, we obtain, as we shall see later, that

$$g_{n,n}(t) = \frac{\left(\frac{t}{2}\right)^{(n-1)/2}}{\pi^{1/2}\Gamma\left(\frac{n}{2}\right)} \quad K_{(n-1)/2}(t).$$

Letting n=1,3,5 we find that $g_{1,1}(t) = \frac{1}{\pi} K_0,(t)$, $g_{2,3}(t)=\frac{t}{\pi}K_1(t)$, and $g_{5,5}(t)=\frac{t^2}{3\pi}K_2(t)$, where $K_n(t)$ is the modified Bessel function of the second kind of order n. See McLachlan [14], and Watson [21].

We shall now find the p.d.f. for cases where $n=m+2$. If we complete the square in 3.1.1 and let $n=m+2$, we get

$$g_{m+2,m}(t) = \frac{(\frac{t}{2})^m}{\Gamma(\frac{m}{2})\Gamma(\frac{m}{2}+1)} \left[\int_1^{\infty} e^{-ty}(y^2-1)^{\frac{m}{2}-1} dy + \int_1^{\infty} e^{-ty}(y^2-1)^{\frac{m}{2}-1} y dy \right]$$

Using the fact that

$$K_n(t) = \frac{\pi^{1/2}(\frac{t}{2})^n}{\Gamma(n+\frac{1}{2})} \int_1^{\infty} e^{-ty}(y^2-1)^{n-\frac{1}{2}} dy, \quad \text{and}$$

$$\frac{d}{dt}\left[t^{-n}K_n(t)\right] = -t^{-n}K_{n+1}(t), \text{ we get that}$$

$$g_{m+2,m}(t) = \frac{(\frac{t}{2})^{\frac{m+1}{2}}}{\pi^{1/2}\Gamma(\frac{m}{2}+1)} \left[K_{(m-1)/2}(t) + K_{(m+1)/2}(t)\right]$$

If we now put $m=1,3$ in the above expression we find that

$$g_{3,1}(t) = \frac{t}{\pi}\left[K_0(t) + K_1(t)\right], \quad \text{and}$$

$$g_{5,3}(t) = \frac{t^2}{3\pi}[K_1(t) + K_2(t)].$$

We can now use 3.1.5 to find $g_{5,1}(t)$, $g_{3,1}(t)$, $g_{1,5}(t)$, and $g_{3,5}(t)$. There are still six values remaining to be found, namely, $g_{1,2}(t)$, $g_{1,4}(t)$, $g_{3,2}(t)$, $g_{3,4}(t)$, $g_{5,2}(t)$ and $g_{5,4}(t)$ which can be expressed in terms of the incomplete gamma function. In fact we find that

$$g_{n,2}(t) = \frac{e^t}{2^{n/2}\Gamma(\frac{n}{2})} \int_{2t}^{\infty} e^{-y} y^{\frac{n}{2}-1}\, dy, \quad \text{and}$$

$$g_{n,4}(t) = \frac{e^t}{\Gamma(\frac{n}{2})2^{\frac{n}{2}+1}} \left[\int_{2t}^{\infty} e^{-y} y^{\frac{n}{2}} dy - 2t \int_{2t}^{\infty} e^{-y} y^{\frac{n}{2}-1} dy \right].$$

Letting n=3,5 we get $g_{3,2}(t)$, $g_{5,2}(t)$, $g_{3,4}(t)$, and $g_{5,4}(t)$. Using 3.1.5 we obtain $g_{1,2}(t)$ and $g_{1,4}(t)$. Now we have all $g_{n,m}(t)$ for $n,m=1,2,3,4,5$, and together with $g_{2,m}(t)$, 3.1.4 and 3.1.5 we can get all the remaining $g_{n,m}(t)$.

Our second objective, then, is to find $\Pr[T_{1,m} > x]$ for all m. We shall first give a method of evaluating this when m

is even, and then a method for any m.

<u>Evaluation of</u> $\Pr[T_{1,m} > x]$, <u>when m is even</u>

$$\Pr[T_{1,m} > x] = \frac{1}{\Gamma(\frac{1}{2})\Gamma(\frac{m}{2})} \int_x^{\infty} e^{-t} H(t)\, dt, \text{ where}$$

$$H(t) = \int_0^{\infty} e^{-2y} y^{\frac{m}{2}-1} (y+t)^{-\frac{1}{2}}\, dy.$$

Expand $H(t)$ in a Maclaurin series about $t=0$. Now

$$H^{(k)}(t) = (-1)^k \pi^{-\frac{1}{2}} \Gamma(k+\tfrac{1}{2}) \int_0^{\infty} e^{-2y} y^{\frac{m}{2}-1} (y+t)^{-(k+\frac{1}{2})} dy,$$

and $$H^{(k)}(0) = (-1)^k \pi^{-\frac{1}{2}} \Gamma(k+\tfrac{1}{2})\, \Gamma(\tfrac{m}{2}-k-\tfrac{1}{2})\, 2^{k+\frac{1}{2}-\frac{m}{2}} .$$

We may write

$$H(t) = \sum_{k=0}^{r} \frac{t^k}{k!} H^{(k)}(0) + R_r(t), \quad \text{where}$$

$$R_r(t) = \frac{1}{r!} \int_0^t (t-y)^r \, H^{(r+1)}(y)\, dy.$$

Note that $\left| H^{(r+1)}(y) \right| \leq \left| H^{(r+1)}(0) \right|$, so that

$$\left| R_r(t) \right| \leq \frac{t^{r+1}}{(r+1)!} \left| H^{(r+1)}(0) \right| .$$

Therefore,

$$\left| \Pr[T_{1,m} > x] - \frac{1}{\Gamma(\frac{1}{2})\Gamma(\frac{m}{2})} \sum_{k=0}^{r} \frac{H^{(k)}(0)}{k!} \int_x^{\infty} e^{-t}\, dt \right| \leq$$

$$\frac{\left| H^{(r+1)}(0) \right|}{\Gamma(\frac{1}{2})\Gamma(\frac{m}{2})(r+1)!} \int_x^{\infty} e^{-t}\, t^{r+1}\, dt.$$

Hence, the error committed by stopping after any term is less in magnitude than the first term neglected.

<u>Evaluation of $\Pr[T_{1,m} > x]$ for any m.</u>

$$\Pr[T_{1,m} > x] = \frac{1}{\pi^{1/2}\Gamma(\frac{m}{2})} \int_x^{\infty} e^{-t} t^{\frac{m-1}{2}} \int_0^{\infty} e^{-2yt} y^{\frac{m}{2}-1} (1+y)^{-\frac{1}{2}}\, dy\, dt.$$

Let $H(t) = (1+t)^{-\frac{1}{2}}$, then

$$H(t) = \pi^{-\frac{1}{2}} \sum_{j=0}^{r} \frac{(-t)^j}{j!} \Gamma(j + \tfrac{1}{2}) + R_r(t), \text{ where}$$

$$R_r(t) = \frac{1}{r!} \int_0^t (t-y)^r H^{(r+1)}(y)\, dy, \text{ and where}$$

$$H^{(r+1)}(y) = (-1)^{r+1} \pi^{-\frac{1}{2}} \Gamma(r+\tfrac{3}{2}) (1+t)^{-(r+\frac{3}{2})}. \text{ Then}$$

Then $\left| H^{(r+1)}(y) \right| \leq \left| H^{(r+1)}(0) \right|$, so that

$$\left| R_r(t) \right| \leq \frac{t^{r+1}}{(r+1)!} \left| H^{(r+1)}(0) \right| .$$

Therefore

$$\left| \Pr[T_{1,m} > x] - \frac{1}{\pi\Gamma(\frac{m}{2})} \sum_{j=0}^{r} \frac{(-1)^j}{j!} \frac{\Gamma(j+\frac{1}{2})\Gamma(\frac{m}{2}+j)}{2^{\frac{m}{2}+j}} I_{j+\frac{1}{2}}(x) \right| \leq$$

$$\frac{\Gamma(r+\frac{3}{2})\Gamma(\frac{m}{2}+r+1)}{\pi\Gamma(\frac{m}{2})(r+1)!\,2^{\frac{m}{2}+r+1}} I_{r+\frac{3}{2}}(x), \text{ where}$$

$$I_{j+\frac{1}{2}}(x) = \int_{x}^{\infty} e^{-t}\, t^{-(j+\frac{1}{2})}\, dt, \qquad j=0,1,\ldots\ .$$

Hence, the error committed is less in magnitude than the first term neglected. We find that

$$I_{j+\frac{1}{2}}(x) = \frac{e^{-x}\, x^{-(j+\frac{1}{2})}}{j-\frac{1}{2}} - \frac{1}{j-\frac{1}{2}}\, I_{j-\frac{1}{2}}(x) \qquad j=0,1,\ldots\ .$$

By repeating the above recurrence relation several times, we find that all $I_{j+\frac{1}{2}}(x)$ can be made to fall back on $\int_{x}^{\infty} e^{-t} t^{1/2}\, dt$, which we can get from a table of the incomplete gamma function.

Briefly summarizing, then, we have shown how to evaluate any $g_{n,m}(t)$ and $Pr[T_{1,m} > x]$ so that we can use 3.1.6 and 3.1.7 to find any c.d.f.

<u>Inequalities</u>.

In what follows, under this heading, we shall discuss certain inequalities related to the distribution of the difference between two independent chi-squares considered in the preceding sub-section, 3.1, of section 3.

<u>If n = 1,</u>

$$g_{1,m}(t) = \frac{e^{-t}t^{-\frac{1}{2}}}{\Gamma(\frac{1}{2})\Gamma(\frac{m}{2})} \int_0^\infty e^{-2y} (1+\frac{y}{t})^{-\frac{1}{2}} y^{\frac{m}{2}-1} dy,$$

and since $e^{-\frac{y}{2t}} \leq (1+\frac{y}{t})^{-\frac{1}{2}} \leq 1$, it follows that

$$3.1.8 \quad 2^{-\frac{m}{2}} h_1(t)(1+\frac{1}{4t})^{-\frac{m}{2}} \leq g_{1,m}(t) \leq 2^{-\frac{m}{2}} h_1(t), \; t > 0,$$

and

$$3.1.9 \quad 2^{-\frac{m}{2}} [1-H_1(x)] \geq \Pr[T_{1,m} > x] \geq 2^{-\frac{m}{2}}(1+\frac{1}{4x})^{-\frac{m}{2}}[1-H_1(x)],$$

where $\quad H_n(x) = \int_0^x h_n(t)\, dt, \qquad x > 0.$

<u>If n,m > 2, then since</u>

$(y+t)^{\frac{n+m}{2}-2} \geq (y+t)^{\frac{n}{2}-1} y^{\frac{m}{2}-1} \geq y^{\frac{n+m}{2}-2}$, it follows that

3.1.10

$$\frac{e^{t}}{2^{\frac{n+m}{2}-1}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}\int_{2t}^{\infty} e^{-y}\, y^{\frac{n+m}{2}-2}\, dy \geq g_{n,m}(t) \geq \frac{e^{-t}\Gamma(\frac{n+m}{2}-1)}{2^{\frac{n+m}{2}-1}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} .$$

Again, since $1 \leq (1+\frac{y}{t})^{\frac{n}{2}-1} \leq e^{\frac{(n-2)y}{2t}}$, it follows from 3.1.1 that,

for $t \geq (\frac{n-2}{4})$,

3.1.11 $2^{-\frac{m}{2}} h_n(t) \leq g_{n,m}(t) \leq 2^{-\frac{m}{2}} h_n(t) \left[1 - \frac{(n-2)}{4t}\right]^{-\frac{m}{2}}$

Now letting $t \longrightarrow \infty$ in 3.1.8 and 3.1.11 we have

3.1.12 $g_{n,m}(t) \sim 2^{-\frac{m}{2}} h_n(t)$, so that

$g_{n,m}(t)$ has the same order of contact at $+\infty$ as the p.d.f. of χ^2 with n degrees of freedom. Similarly, $g_{n,m}(t)$ has the same order of contact at $-\infty$ as the p.d.f. of χ^2 with m degrees of freedom.

In 3.1.10 letting $t \longrightarrow 0$, we have

$$3.1.13 \qquad g_{n,m}(0) = \frac{\Gamma(\frac{n+m}{2} - 1)}{2^{\frac{n+m}{2}-1}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} \quad , \text{ and so}$$

the frequency curve does not meet the origin.

The case n=m.

Under this heading we shall consider certain properties of the p.d.f. and c.d.f. of $T_{n,n}$. In this case the p.d.f. of $T_{n,n}$ is the same as that of the sum of products of pairs of independent $N(0, 1)$ variates.

On the p.d.f. of $T_{n,n}$.

The p.d.f. is symmetric and so we shall consider $g_{n,n}(t)$ for $t \geq 0$. If we put n=m in 3.1.13, 23 get

$$3.1.14 \qquad g_{n,n}(0) = \frac{\Gamma(n-1)}{2^{n-1}\Gamma^2(\frac{n}{2})} = \frac{\Gamma(\frac{n-1}{2})}{2\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} .$$

We can show that $g_{n,n}(0)$ is a decreasing function of n.

Differentiating 3.1.14 we have

$$\frac{2\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \; g'_{n,n}(0) = \frac{\Gamma'(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2})} - \frac{\Gamma'(\frac{n}{2})}{\Gamma(\frac{n}{2})} .$$

From Cramer ([5], p. 131), we have that $\Gamma'(n)/\Gamma(n)$ is an increasing function and so $g_{n,n}(0)$ is a decreasing function of n. That is,

$$\frac{d}{dn}\, g_{n,n}(0) < 0. \tag{3.1.15}$$

Finally, it is easy to show that

$$\frac{d}{dt}\, g_{n,n}(t) = 0 \qquad \text{at } t=0, \text{ and} \tag{3.1.16}$$

consequently that $g_{n,n}(t)$ has its maximum at the origin.

If in 3.1.1 we put n=m, and let $2x = t(y-1)$, we get

$$g_{n,n}(t) = \frac{\left(\frac{t}{2}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})}\, K_{\frac{n-1}{2}}(t) \quad . \tag{3.1.17}$$

If n=1, $g_{1,1}(t) = \frac{1}{\pi} K_0(t)$. It is known that $K_0(t)$ is asymptotic to both axes and has a logarithmic singularity at the origin. Using a well-known expansion for $K_0(t)$, [14], we find that

$$g_{1,1}(t) \sim \frac{1}{\pi} \log \frac{2}{t}\,, \qquad \text{as } t \longrightarrow 0. \tag{3.1.18}$$

The moment generating function of $T_{n,n}$ becomes

$M(\theta) = (1-\theta^2)^{-\frac{n}{2}}$, and so

$$ET_{n,n}^{2s} = \frac{(2s)!}{s!} \frac{\Gamma(\frac{n}{2}+s)}{\Gamma(\frac{n}{2})} \quad .$$

It is worth noting that if we expand $T_{n,n}^{2s}$ and use the known moments of χ^2, we get as a by-product that

$$\sum_{j=0}^{2s} \binom{2s}{j} \Gamma(\tfrac{n}{2}+2s-j)\, \Gamma(\tfrac{n}{2}+j)\, (-1)^j = \frac{(2s)!}{s!} \Gamma(\tfrac{n}{2}+s)\, \Gamma(\tfrac{n}{2}),$$

where $s=0,1,\ldots$ and $n=1,2,\ldots$.

On the c.d.f. of $T_{n,n}$.

For large values of n we may wish to use the normal approximation, since $T_{n,n}$ is asymptotically normal $(0, \sqrt{n})$. It is here appropriate to use a result due to Berry [3]. Let $X=\sum_{i=1}^{n} x_i$, where the x_i are independent r.v.'s. Let $EX=\alpha$, Var $X=\sigma^2$, the c.d.f. of X be $F(x)$, $\lambda(X_i) = E\,|X_i|^3/\text{Var } X_i$, $\Lambda = \max_i \lambda(X_i)$, and

$G(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$, then

$$\sup_{-\infty < x < \infty} \left| F(x) - G\left(\frac{x-\alpha}{\sigma}\right) \right| \leq \frac{C\Lambda}{\sigma}, \text{ where}$$

$$(2\pi)^{-\frac{1}{2}} \leq C \leq 1.88, \text{ according to Berry.}$$

If we let $U_i = \frac{1}{2}(X_i^2 - Y_i^2)$, then we may write $T_{n,n} = \sum_1^n U_i$, where X_i and Y_i are independent $N(0, 1)$ variates. Putting $n=m=1$, and $2x = t(\sec\theta - 1)$ in 3.1.1, we find that $E\left|U_i\right|^3 = 8/\pi$. If $F_n(x)$ is the c.d.f. of $T_{n,n}$, then

$$3.1.19 \qquad \sup_{-\infty < x < \infty} \left| F_n(x\sqrt{n}) - G(x) \right| \leq \frac{8C}{\pi\sqrt{n}}$$

Consider next $\int_x^{\infty} t^n K_n(t)\, dt$. Integrate by parts where

$U = t\,K_n(t)$, $dV = t^{n-1}\,dt$, and use the relation $t\,K_n'(t) = nK_n(t) - t\,K_{n+1}(t)$, getting

$$\int_x^{\infty} t^{n+1}\,K_{n+1}(t)\,dt = x^{n+1}\,K_n(x) + (2n+1)\int_x^{\infty} t^n\,K_n(t)\,dt.$$

Replacing n by $(n-1)/2$, we have

$$3.1.20 \qquad F_{n+2}(x) = F_n(x) - \frac{2\,(x/2)^{\frac{(n+1)}{2}}}{n\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})}\,K_{(n-1)/2}\,(x)$$

$$= F_n(x) - \frac{x}{n}\,g_{n,n}(x)$$

Hence, knowing that $g_{2,2}(x) = \frac{e^{-x}}{2}$, $g_{4,4}(x) = \frac{e^{-x}}{4}\,(1+x)$, $F_2(x) = 1 - \frac{e^{-x}}{2}$, we can obtain all $F_{2n+2}(x)$. Similarly, having a table of the Bessel functions $K_0(x)$, $K_1(x)$, ... we can get all $F_{2n+1}(x)$, if we know $F_1(x)$.

<u>Evaluation of $F_1(x)$.</u>

We have $x[1 - F_1(x)] = \int_x^{\infty} K_0(t)\,dt$. Using tables of

$K_0(t)$ we could evaluate $F_1(x)$ by numerical quadrature; however, we give an alternative method here. Integrate by parts where $U = 1/t$, $dV = tK_0(t)\,dt$, and use the relation $\int t^n K_{n-1}(t)\,dt = -t^n K_n(t)$, $n=1,2,\ldots$, getting

$$\int_x^{\infty} K_0(t)\,dt = K_1(x) - \int_x^{\infty} K_1(t)\,\frac{dt}{t} .$$

If we carry out this integration by parts repeatedly, we find by induction that we get an alternating series, in which the error committed by stopping with a given term is less in magnitude than the first term neglected. If d_r is the r-th term of the series, then

$$d_r = \frac{(-1)^{r+1}\, 2^r\, \Gamma(r+\frac{1}{2})}{\Gamma(\frac{1}{2})\, x^{r-1}}\, K_r(x), \qquad r=1,2,\ldots .$$

Furthermore, if s and k are any two positive integers,

3.1.21 $$\sum_{1}^{2s-1} d_r \geq \pi\,[1 - F_1(x)] \geq \sum_{1}^{2k} d_r .$$

3.2 Special case 2.

In this section we shall give an expression for the p.d.f. of an indefinite q.f. when the latent roots of the matrix of the q.f. are equal in pairs.

Theorem 3.2.

Let

$$Q_n = \frac{1}{2}(a_1 Y_1 + \dots + a_k Y_k - a_{k+1} Y_{k+1} - \dots - a_n Y_n),$$

where the $a_i > 0$, and the Y_i are independent r.v.'s each having a chi-square distribution with two degrees of freedom. If $f(q)$ is the p.d.f. of Q_n, then

$$f(q) = \sum_{j=1}^{k} e^{-\frac{q}{a_j}} a_j^{n-2} \prod_{\substack{i=1 \\ i \neq j}}^{k} (a_j - a_i)^{-1} \prod_{s=k+1}^{n} (a_j + a_s)^{-1}, \quad q > 0$$

$$= \sum_{j=k+1}^{n} e^{\frac{q}{a_j}} a_j^{n-2} \prod_{\substack{i=1 \\ i \neq j}}^{k} (a_j + a_i)^{-1} \prod_{s=k+1}^{n} (a_j - a_s)^{-1}, \quad q < 0 \ .$$

Proof.

The moment generating function of Q_n is

$M(t) = [(1-a_1 t)\dots(1-a_k t)(1+a_{k+1} t)\dots(1+a_n t)]^{-1}$. Then

$$f(q) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itq} M(it)\, dt.$$

We can evaluate $f(q)$ using contour integration. Let us take as our contour in the complex plane, the real line from $-R$ to $+R$ and then the semi-circle of radius R, in the lower half-plane if $q > 0$, and in the upper half-plane when $q < 0$. In both cases the value of the integral around the semi-circle tends to zero as $R \longrightarrow \infty$, if $n \geq 2$. Hence

$$f(q) = 2\pi i\,[\text{Sum of residues of the integrand at } t = -\tfrac{i}{a_j},\ j=1,\dots,k],\quad q > 0, \text{ and}$$

$$= 2\pi i\,[\text{Sum of residues of the integrand at } t = \tfrac{i}{a_j},\ j=k+1,\dots,n],\quad q < 0.$$

Evaluating the residues, we get the form stated in the theorem.

CHAPTER IV

FURTHER BOUNDS ON THE C.D.F. OF A DEFINITE QUADRATIC FORM IN INDEPENDENT IDENTICALLY DISTRIBUTED VARIATES - EACH CENTRAL NORMAL OR MORE GENERAL

4.1 The central normal case.

If we make use of the p.d.f. of Q_{2m} when the latent roots of the matrix of Q_{2m} are equal in pairs, we can obtain some convenient bounds on the p.d.f. of Q_{2n} when the latent roots are not necessarily equal in pairs. Let

$$Q_{2m} = \frac{1}{2}(a_1X_1^2+\ldots+a_mX_m^2+a_{m+1}X_{m+1}^2+\ldots+a_{2m}X_{2m}^2), \text{ where } a_i > 0, \ i=1,\ldots,2m,$$

and where $a_1=a_{m+1}$, $a_2=a_{m+2}$, ..., $a_m=a_{2m}$, then the moment generating function of Q_{2m} is

$$M(t) = [(1-a_1t)\ldots(1-a_mt)]^{-1}, \text{ and hence the p.d.f.}$$

$$\text{of } Q_{2m} \text{ is } h(q) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itq}\, M(it)\, dt.$$

Using the calculus of residues we readily find that

4.1.1 $$h(q) = \sum_{j=1}^{m} e^{-\frac{q}{a_j}} a_j^{m-2} \prod_{\substack{k=1 \\ k \neq j}}^{m} (a_j - a_k)^{-1}$$

Suppose now that we want the p.d.f. of

$Q_{2n} = \frac{1}{2}(a_1X_1^2+\ldots+a_{2n}X_{2n}^2)$, where $a_1 \geq a_2 \geq \cdots \geq a_{2n} > 0$.

We form the two expressions

$Q_U = \frac{1}{2}[a_1(X_1^2+X_2^2) + a_3(X_3^2+X_4^2) + \ldots + a_{2n-1}(X_{2n-1}^2 + X_{2n}^2)]$, and

$Q_L = \frac{1}{2}[a_2(X_1^2+X_2^2) + a_4(X_3^2+X_4^2) + \ldots + a_{2n}(X_{2n-1}^2 + X_{2n}^2)]$, so that

$$Q_U \geq Q_{2n} \geq Q_L \quad .$$

We can find the p.d.f. for Q_U and Q_L by using 4.1 so that we have bounds on the c.d.f. of Q_{2n}. Let $f_U(q)$, $f(q)$, $f_L(q)$ be the p.d.f.'s of Q_U, Q_{2n}, Q_L respectively. Then

$$\int_0^t f_U(q)\, dq \leq \Pr[Q_{2n} \leq t] \leq \int_0^t f_L(q)\, dq, \text{ where}$$

$$f_U(q) = \sum_{j=1}^{n} e^{\frac{q}{a_{2j-1}}} a_{2j-1}^{n-1} \prod_{\substack{k=1 \\ k\neq j}}^{n} (a_{2j-1} - a_{2k-1})^{-1}, \text{ and}$$

$$f_L(q) = \sum_{j=1}^{n} e^{-\frac{q}{a_{2j}}} a_{2j}^{n-1} \prod_{\substack{k=1 \\ k\neq j}}^{n} (a_{2j} - a_{2k})^{-1} \quad .$$

Remarks.

(i) The above inequality was suggested by Professor Hotelling.

(ii) The above method could be extended to cover the case of an indefinite q.f.

4.2 The general case.

In this section we shall discuss briefly a system of inequalities for the distribution of a definite q.f. Let

$$Q_n = \frac{1}{2}(a_1Y_1^2 + \ldots + a_nY_n^2), \text{ where}$$

$$\Pr[Y_i^2 \geq c] = p, \qquad i=1,\ldots,n. \text{ Then}$$

$$\Pr(Y_1^2 < c,\ldots,Y_n^2 < c) = (1-p)^n. \text{ It follows that}$$

$$\Pr\left[Q_n < \frac{c}{2}(a_1 + \dots + a_n)\right] \geq (1-p)^n \quad .$$

Let $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$, getting the following system of inequalities:

$$\Pr\left[Q_n < \frac{c}{2}\, a_n\right] \leq (1-p)^n \ ,$$

$$\Pr\left[Q_n < \frac{c}{2}(a_{n-1} + a_n)\right] \leq (1-p)^n + \binom{n}{1}(1-p)^{n-1}p,$$

$$\Pr\left[Q_n < \frac{c}{2}(a_{n-2} + a_{n-1} + a_n)\right] \leq (1-p)^n + \binom{n}{1}(1-p)^{n-1}p + \binom{n}{2}(1-p)^{n-2}p^2,$$

$$\dots\dots\dots$$

$$\Pr\left[Q_n < \frac{c}{2}(a_1 + \dots + a_n)\right] \leq 1-p^n$$

The above system was suggested to the writer by Professor Harold Hotelling. The following improvements are due to Professor S. N. Roy.

Suppose that the distribution of Y_1^2 is known and that

$$\Pr(Y_1^2 \geq \frac{2c}{a_n}) = P_1,$$

$$\Pr(Y_1^2 \geq \frac{2c}{a_{n-1}+a_n}) = P_2,$$

$$\Pr(Y_1^2 \geq \frac{2c}{a_{n-2}+a_{n-1}+a_n}) = p_3,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\Pr(Y_1^2 \geq \frac{2c}{a_1+\ldots+a_n} = p_n.$$

Then it follows that:

$$(1-p_n)^n \leq \Pr(Q_n < c) \leq (1-p_1)^n,$$

$$(1-p_n)^n \leq \Pr(Q_n < c) \leq (1-p_2)^n + \binom{n}{1}(1-p_2)^{n-1}\, p_2,$$

$$(1-p_n)^n \leq \Pr(Q_n < c) \leq (1-p_3)^n + \binom{n}{1}(1-p_3)^{n-1} p_3 + \binom{n}{2}(1-p_3)^{n-2}\, p_3^2,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$(1-p_n)^n \leq \Pr(Q_n < c) \leq 1-p_n^n\ .$$

Hence we have one lower bound and a whole system of upper bounds. Obviously, we would want the least upper bound, in practice. Incidentally, $p_1 \leq p_2 \leq \cdots \leq p_n$.

BIBLIOGRAPHY

[1] Anderson, R. L., "Distribution of the Serial Correlation Coefficient", Annals of Mathematical Statistics, XIII (1942), 1-13.

[2] Aroian, L. A., "The Probability Function of the Product of two Normally Distributed Variables", Annals of Mathematical Statistics, XVIII (1947), 265-271.

[3] Berry, A. C., "The Accuracy of the Gaussian Approximation to the Sum of Independent Variates", Transactions of the American Mathematical Society, XLIX (1941), 122-136.

[4] Craig, C. C., "On the Frequency Function of XY", Annals of Mathematical Statistics, VII (1936), 1-15.

[5] Cramér, H., Mathematical Methods of Statistics, Princeton, University Press, 1946.

[6] Durbin, J., Watson, G. S., "Testing for Serial Correlation in Least Squares Regression", Biometrika, XXXVII (1950), 409, 428.

[7] Edwards, J., A Treatise on the Integral Calculus, Vol. II, London, Macmillan and Co., 1921-22.

[8] Gurland, J., "Distribution of Quadratic Forms and Ratios of Quadratic Forms", Annals of Mathematical Statistics, To be published.

[9] Hotelling, H., "The Selection of Variates for Use in Prediction with Some Comments on the Problem of Nuisance Parameters", Annals of Mathematical Statistics, XI (1940), 271-283.

[10] Hotelling, H., "Some New Methods for Distributions of Quadratic Forms", abstract, Annals of Mathematical Statistics, XIX (1948), 119.

[11] Kendall, M. G., The Advanced Theory of Statistics, Vol. I, London, Charles Griffin and Company Limited, 1947.

[12] Koopmans, T., "Serial Correlation and Quadratic Forms in Normal Variables", Annals of Mathematical Statistics, XIII, (1942), 14-33.

[13] McCarthy, M. D., "On the Application of the Z-Test to Randomized Blocks", Annals of Mathematical Statistics, X (1939), 337.

[14] McLachlan, N. W., Bessel Functions for Engineers, Oxford, University Press, 1934.

[15] Pearson, K., Stouffer, S.A. David, F.N. "Further Application in Statistics of the $T_m(x)$ Bessel Function", Biometrika, XXIV (1932), 293-350.

[16] Robbins, H. E., "The Distribution of a Definite Quadratic Form", Annals of Mathematical Statistics, XIX (1948), 266-270.

[17] Robbins, H. E., Pitman, E. J. G., "Application of the Method of Mixtures to Quadratic Forms in Normal Variates", Annals of Mathematical Statistics, XX (1949), 552-560.

[18] Szegö, G., Orthogonal Polynomials, American Mathematical Society Colloquium Publications, Vol. XXIII, N.Y. 1939.

[19] von Neumann, J., Bellinson, H.R., Kent, R.H. Hart, B.I. "The Mean Square Successive Difference", Annals of Mathematical Statistics XII (1941), 153-162.

[20] von Neumann, J., "Distribution of the Ratio of the Mean Square Successive Difference to the Variance", Annals of Mathematical Statistics (1941), 367-395.

[21] Watson, G. N., Theory of Bessel Functions, Cambridge, University Press, 1945.